# Calculation and Analysis of Low Frequency Normal Modes for DNA

**TAP HA DUONG, KRYSTYNA ZAKRZEWSKA**

*Laboratoire de Biochimie Théorique, URA 77, CNRS, Institut de Biologie Physico-Chimique, 13, rue Pierre et Marie Curie, 75005 Paris, France*

*Received 8 May 1996; accepted 29 August 1996*

**ABSTRACT:** Normal mode calculations for two alternating sequence dodecamers in A, B, and Z conformations have been performed in dihedral angle space extended to endocyclic valence angles to account for sugar ring flexibility. Normal modes are analyzed in terms of helicoidal and backbone parameter variations with special attention being paid to global deformations of the double helix such as bending, twisting, or stretching. Results show that the allomorphic form of DNA has the largest influence on the flexibility of the sugar-phosphate backbone. Amplitudes of these vibrations follow the order: B > Z > A. In contrast, the amplitudes of helicoidal parameter variations are much more dependent on the base sequence. Global deformations of the double helix occur with characteristic times in the range of 1 to 10 ps and can be of mixed character, the strongest bending mode being at the same the time strongest stretching mode. © 1997 by John Wiley & Sons, Inc. *J Comput Chem* **18**: 796–811, 1997

**Keywords:** DNA; normal mode; flexibility; modeling; DNA bending

## Introduction

Low frequency internal motions of macromolecules are known to play an important role in biological processes such as enzymatic catalysis and molecular recognition.[1–3] However, these large amplitude motions, with characteristic time scales of 1 ps or more, are the most difficult to study experimentally and often produce results that are difficult to analyze in conformational terms.[4–7] We are particularly interested in the role such motions may play in the interactions between proteins and nucleic acids, the specificity of which is vital to the functioning of the genetic apparatus. Over recent years, structural information on protein–DNA complexes has increased rapidly, but there is still no general understanding of the mechanics of such interactions. The early notion that specificity is linked to ''direct read out'' via hydrogen bonding between the amino acid side chains and nucleic acid bases has been shown to be insuf-

*Correspondence to*: T. Ha Duong

Contract grant sponsor: Association for International Research

 CCC 0192-8651 / 97 / 060796-16

ficient. A major role is now attributed to ''indirect read out,'' which includes both a static part, the fine structure of DNA (including its the environment of water and counterions), and a dynamic part linked to sequence dependent flexibility.

Although molecular dynamics (MD) is now a standard technique for studying the dynamics of biological macromolecules, it is computationally expensive, particularly when applied to simulating slow movements. An alternative approach in this case is provided by normal mode (NM) analysis which can describe the vibrations around a chosen energy minimum rapidly, even for very large systems, by making an harmonic approximation to the local energy hypersurface.[8,9] Most calculations in this area, however, have been carried out for proteins and have also used a standard Cartesian coordinate space (CCS) to describe the macromolecule.

Based on recent results of Sunada and Go[10] which show that the harmonicity of conformational energy hypersurface can be assumed over a wider range in dihedral angle space (DAS) than in CCS, we have developed a program that calculates normal modes for nucleic acids in DAS, including an extension for the treatment of sugar ring flexibility. This approach is advantageous for large systems since it leads to an important reduction in the number of variables, although, as recent studies on deca-alanine by Kitao and Go[11,12] indicate, they maintain an almost perfect one-to-one correspondence with the full CCS modes in the low frequency range (differences being limited to a uniform frequency shift below 128 $cm^{-1}$).

In addition to developing this program, we have developed techniques for analyzing the collective deformations occurring within the double helix. A normal mode calculation yields a set of coefficients for each mode over all physical variables which, especially for the lowest frequencies, are very difficult to interpret. To understand these movements we have built a series of structures along the direction of each mode and then analyzed these structures with the CURVES program.[13] This algorithm provides a rigorous helicoidal description of any irregular nucleic acid fragment. It yields both helicoidal parameters and the overall locus of the global helical axis. This procedure allows us to follow variations in geometrical parameters in an automatic and objective way, and to detect movements localized on specific base pairs (e.g., opening), between base pairs (e.g., roll), or of a global nature such as the bending, twisting, and stretching of the double helix.

## Methodology

### NORMAL MODE ANALYSIS

Normal mode theory provides an analytical solution of the equations of motion, by approximating the potential energy of the system as a quadratic function of its $n$ coordinates, $q_i$, around the potential energy minimum $q_i^0$. The theory shows that a solution of the Lagrange equations can be found in the form:[14]

$$q_i(t) = q_i^0 + \sum_{k=1}^{n} A_{ik}\alpha_k \cos(\omega_k t + \delta_k)$$

where $\alpha_k$ and $\delta_k$ depend on initial conditions. The unknowns, $A_{ik}$ and $\omega_k$, are given by solving the matrix equation:

$$\mathbf{HAW} = \mathbf{FA}$$

where **W** is a diagonal positive matrix whose elements are:

$$W_{ik} = \delta_{ik}\omega_k^2$$

Elements of the potential energy Hessian matrix, **F**, are written:

$$F_{ij} = \left(\frac{\partial Ep}{\partial q_i\, \partial q_j}\right)_0$$

while elements of the kinetic energy matrix, **H**, are given by:

$$H_{ij} = \sum_{a=1}^{N} m_a \left(\frac{\partial \mathbf{r}_a}{\partial q_i}\right)_0 \left(\frac{\partial \mathbf{r}_a}{\partial q_j}\right)_0$$

### THERMODYNAMICS OF VIBRATION

The contribution of harmonic vibrations to thermodynamic properties can be calculated using the following equations:

$$Z_{\text{vib}} = \prod_{k=1}^{M} \frac{e^{-\hbar\omega_k/2KT}}{1 - e^{-\hbar\omega_k/KT}}$$

$$F_{\text{vib}} = -KT \log Z_{\text{vib}}$$

$$S_{\text{vib}} = -\frac{\partial F_{\text{vib}}}{\partial T}$$

where $Z_{\text{vib}}$, $F_{\text{vib}}$, and $S_{\text{vib}}$ represent, respectively, the partition function, the Gibbs free energy, and the entropy of vibration. These calculations can provide information on the vibrational properties of different DNA fragments.

## DNA MODELING IN DAS

Modeling the flexibility of DNA requires taking into account the flexibility of its sugars. This in turn requires extending a pure dihedral angle representation to include the endo valence angles of the sugar, without which ring repuckering cannot occur (all bond lengths effectively being frozen). Ring closure conditions imply that the dihedral and valence angles of the sugar are coupled by complicated nonlinear equations which reduce 10 physical variables to 4 true degrees of freedom. We treat this problem using the approach developed by Lavery et al.[15] which temporarily considers a ring as a linear molecule (by breaking one ring bond), and enforces closure using a quadratic distance constraint. The variables which then describe the sugar conformation are two dihedral angles and three valence angles as shown in Figure 1. In addition, it is assumed that when an endo-valence angle increases by $x$ then the exo-valence angles made by any atoms joined to the ring at this point with the ring bonds will decrease by $x/2$, maintaining the symmetric position of the substituents.

Last, to position the two strands of the DNA duplex with respect to one another, we use three atoms in each strand related by six variables, a distance, two valence angles, and three dihedral angles, as shown in Figure 2. Such a formulation allows us to use the same formulas in the calculation of derivatives as for the internal variables.

## FORCE FIELD

The function used to calculate the potential energy of a nucleic acid and its derivatives is the FLEX force field developed by Lavery et al.[16] and

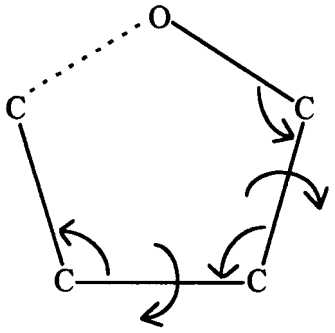


**FIGURE 1.** Definition of the five independent variables of the sugar ring. The dotted line corresponds to the bond treated as a distance constraint.

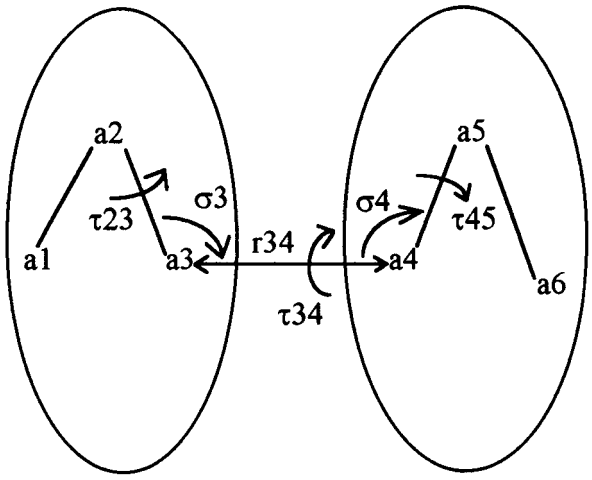


**FIGURE 2.** Description of the six variables defining the position of the second strand with respect to the first.

is given by:

$$E_p = \sum_i V_{i(\mathrm{val})}(\sigma_i - \sigma_i^0)^2 + \sum_i \frac{V_{i(\mathrm{tor})}}{2}(1 \pm \cos(n_i\tau_i)) + \sum_{a,b}\left[\frac{q_a q_b}{\varepsilon r_{ab}} + \left(-\frac{A_{ab}}{r_{ab}^6} + \frac{B_{ab}}{r_{ab}^{12}}\right)\right] + \sum_{HB}\left[\cos\theta\left(-\frac{A_{ab}^{HB}}{r_{ab}^6} + \frac{B_{ab}^{HB}}{r_{ab}^{12}}\right) + (1-\cos\theta)\left(-\frac{A_{ab}}{r_{ab}^6} + \frac{B_{ab}}{r_{ab}^{12}}\right)\right] + \sum_{\mathrm{rings}} K(r_{12} - r_{12}^0)^2$$

The first two terms represent, respectively, the energy of deformation of valence and torsion angles. The third and fourth terms represent the nonbonded interactions (electrostatic, Lennard–Jones, and hydrogen bond) between all pairs of atoms separated by at least three chemical bonds. The dielectric function, $\varepsilon$, is a sigmoidal function of the distance separating two charges, whose slope and plateau can be controlled.

$$\varepsilon(r_{ab}) = D - \frac{D-1}{2}\left[(Cr_{ab})^2 + 2Cr_{ab} + 2\right] \times \exp(-Cr_{ab})$$

In this study, the slope value is fixed to 0.356 and the plateau value to 78. The net phosphate charges were reduced to $-0.5$e.[17,18] The last term in the FLEX energy function takes into account the penalty forces introduced for the ring closure conditions.

## CALCULATION OF MATRICES F AND H

Elements of the potential energy matrix, **F**, are calculated analytically. We first calculate atomic forces on all the atoms and then, appropriately, combine these forces over the moving atoms for each independent variable of the system taking into account dependent variable moves associated with the sugar rings.[16] The second derivatives of the potential energy depend on atomic positions, atomic forces, and their derivatives. Access to full mathematical expressions for second derivatives of Ep is described in Appendix A.

Concerning the kinetic matrix, **H**, special attention must be paid to the derivatives of atomic position vector $\mathbf{r}_a$ with respect to independent variables $q_i$. Because the internal kinetic energy must not include overall translation and rotation of the system, the change $\delta\mathbf{r}_a$ caused by $\delta q_i$ must not move or rotate the center of mass. An elegant analytical calculation of **H** in DAS has been published by Noguti and Go.[19] They considered the system as two rigid bodies connected by a chemical bond which rotates when the dihedral angle $q_i$ changes (Fig. 3). For each atom $b$ in the first body and each atom $c$ in the second, $\delta\mathbf{r}_b$ and the $\delta\mathbf{r}_c$ are expressed using Eckart's conditions[20] to separate external and internal motion. These can be formulated as:

$$\sum_a m_a \delta\mathbf{r}_a = 0$$

$$\sum_a m_a \mathbf{r}_a^0 \times \delta\mathbf{r}_a = 0$$

where $\mathbf{r}_a^0$ are atomic positions at the energy minimum.

This approach can easily be generalized to the case where the variable $q_i$ is the interstrand distance or one of the two interstrand valence angles. However, for an endocyclic ring valence angle, we have extended the Noguti and Go approach to a molecule regarded as three rigid bodies connected by two chemical bonds (Fig. 4). The third body includes all exocyclic substituent atoms of the ring valence angle. Within this model, when $q_i$ changes its value by $\delta q_i$, each atom $b$ of the first body and each atom $c$ of the second move in such a way that:

$$\delta\mathbf{r}_b = \delta\mathbf{r}_i + \mathbf{w}_1 \times (\mathbf{r}_b^0 - \mathbf{r}_i^0)$$

$$\delta\mathbf{r}_c = \delta\mathbf{r}_i + \mathbf{w}_2 \times (\mathbf{r}_c^0 - \mathbf{r}_i^0)$$

where $\mathbf{w}_1$ and $\mathbf{w}_2$ are the two rotation vectors of the first two bodies, and are related by the expression:

$$\mathbf{w}_2 - \mathbf{w}_1 = \delta q_i \mathbf{e}_i$$

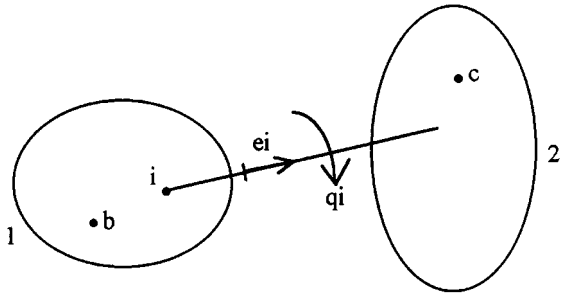


**FIGURE 3.** Molecule treated as two rigid bodies for the calculation of atomic position derivatives with respect to dihedral angles.

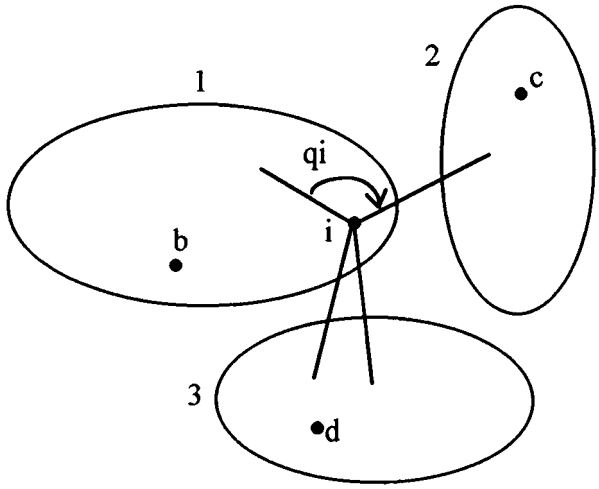


**FIGURE 4.** Molecule considered as three rigid bodies for the calculation of atomic position derivatives with respect to endocyclic valence angles.

Because the third body rotates in such a way as to maintain its symmetric position, the motion of each of its atoms has to satisfy the following two equations:

$$\delta \mathbf{r}_d = \delta \mathbf{r}_i + \mathbf{w}_3 \times (\mathbf{r}_d^0 - \mathbf{r}_i^0)$$
$$\mathbf{w}_3 = \frac{\mathbf{w}_1 + \mathbf{w}_2}{2}$$

By applying Eckart's conditions to $\delta \mathbf{r}_b$, $\delta \mathbf{r}_c$, and $\delta \mathbf{r}_d$, it is therefore possible to analytically formulate the derivatives of atomic positions $\delta \mathbf{r}_a / \delta q_i$ with respect to endocyclic valence angles and to exclude external motion. For more details on the complete expressions of $\delta \mathbf{r}_a / \delta q_i$ see Appendix B.

The analytical formulas derived for matrices **H** and **F** were tested numerically. For matrix **F** analytic first derivatives were used, while in the case of kinetic energy we followed the procedure of Levitt et al.[21] Both numerical and analytical calculation times scale as $n^3$, where $n$ is the number of variables. Our tests, however, show that the analytical calculations are roughly twice as fast as the numerical approach.

### CALCULATION PROTOCOL

The DNA sequences studied were initially constructed and optimized with the JUMNA program[17,18] using helical symmetry constraints to aid the search for global energy minima.[22] These conformations were then reoptimized in the LIGAND program as unconstrained oligomers. Most calculations were carried out with a length of 12 base pairs; however, to elucidate the influence of length on their vibrational properties, and to compare with earlier studies, we also performed the calculations for the oligomers with lengths increasing from 6 to 18 base pairs. LIGAND, derived from an earlier program, CINFLEX,[15] which works in dihedral angle space (including the flexible ring approach described previously), was extended to allow normal mode calculations. After minimization using a modified Newton minimizer (Harwell VA13A), all the vibrational frequencies obtained were positive. For each normal mode a series of structures were created within the limits of thermal energy (0.6 kcal/mol at 300 K) and analyzed by CURVES.[13] The results of this analysis allowed us to define the ''bending mode'' as the mode with the largest range of the angle between the terminal helical axis segments. In the same way the twisting and stretching modes were defined, respectively, as those with the largest range of total twist and rise. This implies that a single mode may have a multiple character, and this indeed occurs as will become clear in the Results section.

## Results

To analyze the influence of allomorphic conformation and of base pair sequence on the dynamic properties of the DNA double helix, we calculated the normal modes for the three canonical DNA forms, A, B, and Z, using the two alternating base sequences, poly(dC-dG) · poly(dC-dG) and poly(dT-dA) · poly(dT-dA). These six resulting structures are termed A-CG, B-CG, Z-CG, A-TA, B-TA, and Z-TA in what follows.

### HISTOGRAM OF FREQUENCIES

In the DNA model used, there are 260 modes for the CG sequences and 272 for the TA sequences (this increase being due to presence of thymine methyl rotations in the latter case). For both sequences, there are 72 modes with frequencies above 450 $cm^{-1}$ (i.e., with periodicities of less than 75 fs), which essentially involve endocyclic valence angles of the sugar. By contrast, all dihedral angle movements have frequencies below 250 $cm^{-1}$ (i.e., periodicities greater than 130 fs). Figure 5 shows the histograms of these lowest frequencies for our six dodecamers. This range of frequencies agrees well with frequencies calculated on a 2′-deoxyguanosine nucleotide by a Fourier transform of a 1-ns molecular dynamic simulation.[23] (Note that the relationship between periodicity, $T$, [ps] and frequency, $\omega$, [$cm^{-1}$] is approximately $T = 33.4/\omega$.) In the part of the spectrum below 250 $cm^{-1}$, we can observe one peak around 60 $cm^{-1}$ for A-DNA, whereas there are two principal maxima for the two other allomorphs occurring, around 20 $cm^{-1}$ and 100 $cm^{-1}$, respectively, in B-DNA, and around 30 $cm^{-1}$ and 80 $cm^{-1}$ in Z-DNA.

### SUGAR PSEUDOROTATIONAL VIBRATION

A very convenient method for describing sugar conformation is the so-called pseudorotational model, in which ring puckering is described by two angular parameters, phase and amplitude, which can be calculated from the ring dihedral angles.[24] Analyzing our spectra in this way we note a weak amplitude of phase vibration in A-

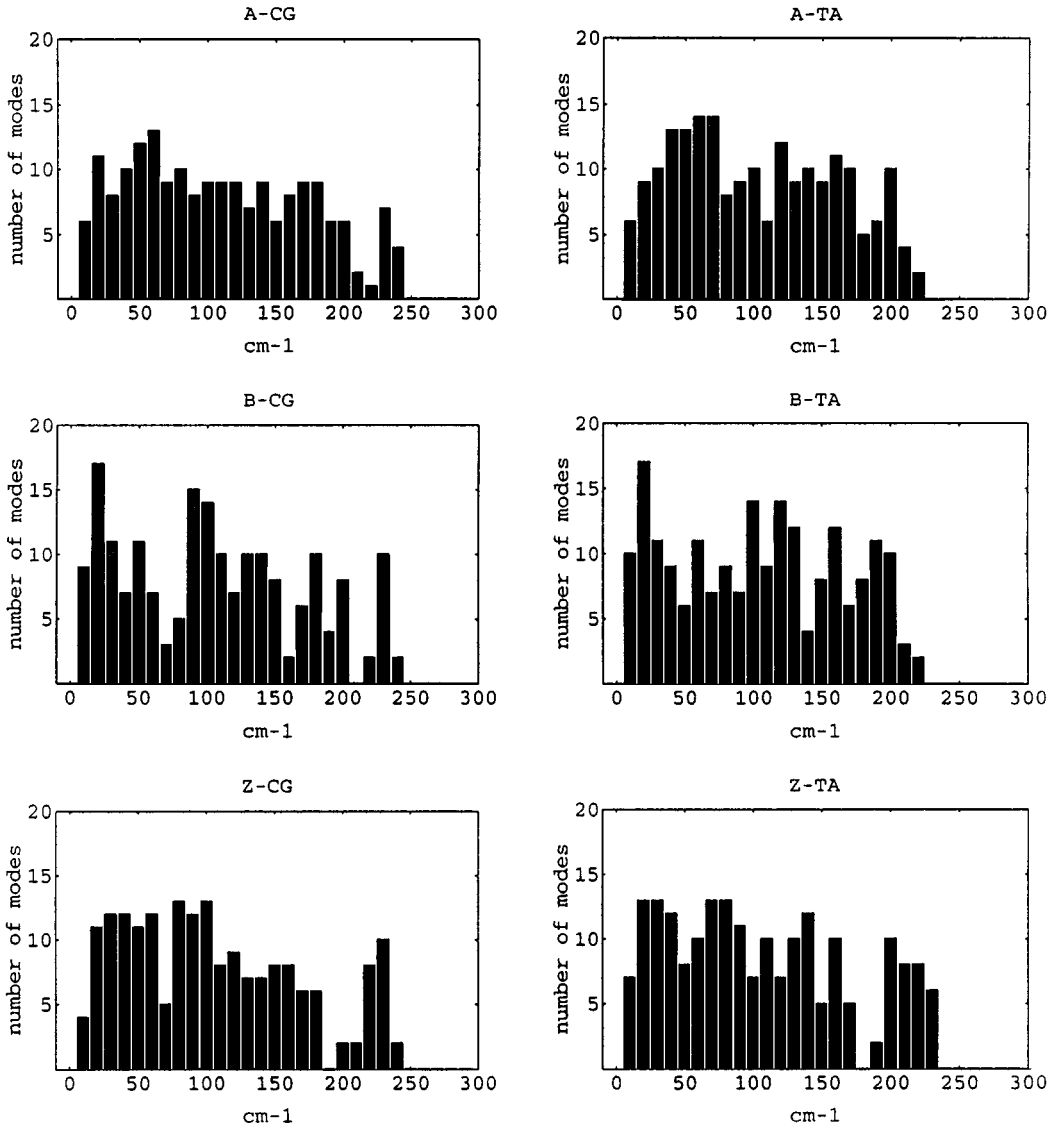


**FIGURE 5.** Vibrational frequency histogram for the dodecamers A-CG, B-CG, Z-CG, A-TA, B-TA, and Z-TA (bar width 10 $cm^{-1}$).

DNA ($< 8°$) occurring between 30 and 80 $cm^{-1}$. In B-DNA the main phase changes occur at 20 $cm^{-1}$ ($> 10°$) with a shoulder between 80 and 120 $cm^{-1}$ ($\sim 6°$). In the case of Z-DNA, there are two distinct peaks at 20 and 80 $cm^{-1}$ ($\sim 10°$) associated with pyrimidines (C2′-endo) and two peaks at 50 and 220 $cm^{-1}$ ($\sim 8°$) for purines (C3′-endo).

In regard to sugar amplitude variations, we observe one peak around 70 to 90 $cm^{-1}$ in A-DNA with a height of $\sim 3°$. In the B-DNA spectra, we observe two peaks, one at 20 $cm^{-1}$ ($\sim 5°$) and the second between 130 and 200 $cm^{-1}$ ($\sim 4°$). In Z-DNA, the pyrimidines are associated with three peaks having roughly the same height ($\sim 3°$) at 50 $cm^{-1}$, 90 $cm^{-1}$, and 130 $cm^{-1}$, whereas purines present two peaks at 70 $cm^{-1}$ ($\sim 4°$) and 220 $cm^{-1}$ ($\sim 3°$). Overall, this analysis provides the same image of sugar flexibility as that involving the valence angle variations; B-DNA sugars are more flexible than A-DNA sugars, whereas the frequencies associated with Z-DNA show mixed behavior.

It is interesting to note that, although phase and amplitude variations along each mode appear to be coupled for each sugar, the nature of this coupling varies dramatically from mode to mode and from residue to residue. Consequently, it seems important to conserve the independence of these two variables for any realistic model of DNA.

### AMPLITUDE OF VIBRATION OF BACKBONE TORSION ANGLES

To understand the vibrational behavior of the DNA backbones we have calculated the range covered by each dihedral angle at 300 K (covering frequencies up to 250 $cm^{-1}$). Table I gives these values for the two central nucleotides of the two sequences studied in each allomorphic form.

Overall, A-DNA again seems to be the least flexible form. The largest variations occur for angles $\alpha$ (P—O5′) and $\gamma$ (C5—C4′) at around 210 $cm^{-1}$ and there are no significant variations below 100 $cm^{-1}$. In contrast, for B-DNA, the $\zeta$ (O3′—P) angle shows a large amplitude of vibration around 20 $cm^{-1}$, and a smaller but still significant peak around 190 $cm^{-1}$. These two peaks also occur in the spectra for the coupled variation $\varepsilon$–$\zeta$ with ranges of 25° and 14°, respectively. This movement represents a flexible crankshaft which displaces the phosphate group and corresponds to the experimentally observed $B_I$–$B_{II}$ transition.[25] Finally, for Z-DNA backbone angles, the larger vibrations concern the $\alpha$ and $\beta$ (O5′—C5′) dihedrals of the pyrimidine–purine steps. These variations occur between 10 and 80 $cm^{-1}$ with a range of around 15° for $\beta$ and 10° for $\alpha$. In the same way as in B-DNA $\varepsilon$–$\zeta$ spectra, the Z-DNA $\alpha$–$\beta$ spectra for pyrimidine–purine steps present an important peak ($\sim$ 30°) around 70 to 80 $cm^{-1}$, suggesting relatively high flexibility. In contrast, the purine–pyrimidine steps seem to be more rigid, with the largest variations involving $\alpha$ and $\gamma$ angles at 220 $cm^{-1}$ (range $\sim$ 8°), similar to those in A-DNA.

### AMPLITUDE OF VIBRATION OF HELICOIDAL PARAMETERS

In Table II we report the maximal ranges of chosen base-pair and inter–base-pair helicoidal parameters. Overall, angular ranges are between 3° and 9° with the largest range variations for the base-pair parameters buckle and propeller. Important variations in translation are restricted to in-

**TABLE I.**
**Maximal Ranges of Backbone Dihedral Angles (Degrees) for the Two Central Base Pairs of the Six DNA Dodecamers Studied.**

| | A-DNA | | | | B-DNA | | | | Z-DNA | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | CpG | GpC | TpA | ApT | CpG | GpC | TpA | ApT | CpG | GpC | TpA | ApT |
| $\alpha$ | 8.5 | 9.1 | 6.0 | 5.9 | 5.9 | 7.6 | 5.1 | 6.8 | 9.9 | 8.5 | 11.8 | 5.8 |
| $\beta$ | 4.0 | 4.0 | 4.0 | 4.5 | 6.2 | 5.6 | 5.9 | 6.2 | 13.3 | 3.9 | 14.0 | 6.1 |
| $\gamma$ | 7.0 | 5.7 | 4.9 | 4.6 | 5.0 | 4.8 | 4.2 | 5.5 | 5.3 | 6.5 | 5.3 | 8.0 |
| $\varepsilon$ | 4.4 | 4.8 | 4.0 | 4.0 | 5.7 | 6.1 | 4.2 | 4.7 | 7.8 | 3.3 | 5.9 | 5.1 |
| $\zeta$ | 3.5 | 3.7 | 3.6 | 3.4 | 19.0 | 19.2 | 9.6 | 13.2 | 5.9 | 3.6 | 5.7 | 4.9 |
| $\chi$ | 7.3 | 6.3 | 4.3 | 3.3 | 8.1 | 8.6 | 5.5 | 3.8 | 3.3 | 5.5 | 6.2 | 7.2 |

**TABLE II.**
**Maximal Ranges of Helicoidal Parameters for the Two Central Base Pairs of the Six Dodecamers Studied (Shift, Slide, Rise, and Stretch in Angstroms, Other Parameters in Degrees).**

| | A-DNA | | | | B-DNA | | | | Z-DNA | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | CpG | GpC | TpA | ApT | CpG | GpC | TpA | ApT | CpG | GpC | TpA | ApT |
| Shift | 0.2 | 0.2 | 0.3 | 0.4 | 0.9 | 1.0 | 0.7 | 0.6 | 0.2 | 0.4 | 0.3 | 0.4 |
| Slide | 0.3 | 0.3 | 0.3 | 0.2 | 0.4 | 0.6 | 0.3 | 0.5 | 0.4 | 0.4 | 0.5 | 0.5 |
| Rise | 0.8 | 0.7 | 0.5 | 0.5 | 0.5 | 0.6 | 0.4 | 0.4 | 0.4 | 0.3 | 1.0 | 0.4 |
| Stretch | 0.3 | 0.2 | 0.2 | 0.2 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.1 | 0.2 | 0.2 |
| Tilt | 3.0 | 2.5 | 2.4 | 2.7 | 2.1 | 3.0 | 2.3 | 2.0 | 3.1 | 4.3 | 2.8 | 3.1 |
| Roll | 4.9 | 4.5 | 4.4 | 4.3 | 6.7 | 5.2 | 5.3 | 4.4 | 3.2 | 4.0 | 3.1 | 4.9 |
| Twist | 3.4 | 3.1 | 2.1 | 2.8 | 2.6 | 3.9 | 2.5 | 2.9 | 2.3 | 2.1 | 3.5 | 2.3 |
| Buckle | 8.4 | 8.4 | 8.4 | 8.4 | 9.9 | 9.7 | 7.9 | 8.0 | 5.0 | 5.2 | 8.1 | 7.9 |
| Propeller | 7.2 | 7.2 | 6.3 | 6.3 | 8.0 | 8.0 | 6.1 | 6.1 | 4.9 | 4.8 | 7.0 | 7.0 |
| Opening | 2.3 | 1.3 | 3.0 | 2.6 | 1.5 | 1.6 | 2.0 | 2.0 | 0.8 | 1.3 | 2.5 | 2.5 |

ter–base-pair rise, which is linked to the overall stretching of the duplex, and to shift and slide in the case of B allomorph. These variations reach 1 Å in contrast to base-pair stretching which is limited to variations of around 0.2 Å (only very small variations are also seen in the variables linking together the two strands of the duplex).

It is interesting to note that a preference for inter–base-pair roll rather than tilt, well known for B-DNA,[26–29] is also seen in this dynamic analysis. However, base sequence effects, such as the differences between purine–pyrimidine and pyrimidine–purine steps, observed in crystallographic structures, do not appear to have a clear dynamic counterpart. It is also worth noting that the range covered by the twist angles in B-DNA (roughly 3°) are well below the range observed for optimal conformations as a function of sequence (roughly 15°). This suggests that vibrations at room temperature will not render such differences undetectable during DNA recognition processes.

## GLOBAL DEFORMATIONS

From a biological point of view the most interesting modes are those which involve global deformations of the double helix such as twisting, bending, and stretching, because such deformations are often induced by protein–DNA interactions. Using CURVES as an analysis tool, we were able to locate three bending modes and one twisting mode among the lowest frequency modes of each of the DNA fragments under study (under 10 $cm^{-1}$; i.e., with characteristic times greater than 3 ps). These modes are however not necessarily of a single character—on the contrary, in most cases they show mixed behavior. Table III summarizes the maximal ranges and frequencies of these global deformations. By looking at the frequencies in the first column of the table (A-CG), we can see that the first twisting mode is, at the same time, the first bending mode and the first stretching mode. The same triple character is found for the mode at 3.7 $cm^{-1}$, but not for the bending mode of 2.1 $cm^{-1}$, which may be considered as pure bending.

Figure 6 shows the variations of helicoidal parameters along the oligomers studied for the two lowest frequency bending modes. We can clearly observe that these global deformations involve the motion of all nucleotides in a cooperative way. Note that the mechanics of bending involves principally rise and roll variations, linked in some cases to changes in twist, but shows little variation in tilt. Looking in more detail at Figure 6a, this explains why for the first bending mode for the A and B conformers is, effectively, also the first stretching mode. Similarly, it is also the first twisting mode for the TA sequence. In Figure 6b the striking impression is that the first overtone of the fundamental bending mode is seen. Here the two halves of the oligomer vibrate in opposite directions resulting in the negligible changes in total twist and rise, thus explaining the absence of this mode in Table III for twisting or stretching deformations. Although the details of these collective modes depend on the allomorphic form and on the base sequence, the general impression is that low frequencies correspond to standing waves along a vibrating string. (The situation is clearly more complicated for the Z conformer due to its more complex dinucleotide repeat structure).

Finally, to illustrate how bending is coupled to roll and twist, we show the time evolution of rise,

**TABLE III.**
**Ranges of Global Twisting, Bending, and Stretching Modes (Stretch in Angstroms, Twist and Bend in Degrees, and Frequencies [Given in Parentheses] in Centimeter $^{-1}$).**

| | A-DNA | | B-DNA | | Z-DNA | |
|---|---|---|---|---|---|---|
| | CG | TA | CG | TA | CG | TA |
| Twist | 19(1.5) | 10(2.1) | 17(3.8) | 10(3.8) | 15(5.6) | 27(2.4) |
| | 16(3.7) | 9(4.9) | 6(8.5) | 10(5.0) | 5(11.0) | 15(2.8) |
| Bend | 10(1.5) | 10(2.1) | 21(1.7) | 15(2.0) | 8(3.1) | 17(2.8) |
| | 12(2.1) | 11(2.7) | 12(2.0) | 12(2.3) | 9(3.8) | 12(5.3) |
| | 8(3.7) | 8(4.9) | 13(3.2) | 11(3.8) | 7(7.2) | 13(12.3) |
| Stretch | 4.0(1.5) | 2.5(2.1) | 2.3(1.7) | 1.6(2.0) | 0.9(3.1) | 3.8(2.4) |
| | 2.2(3.7) | 1.7(4.9) | 2.6(3.2) | 1.4(3.8) | 1.2(5.6) | 1.6(2.8) |

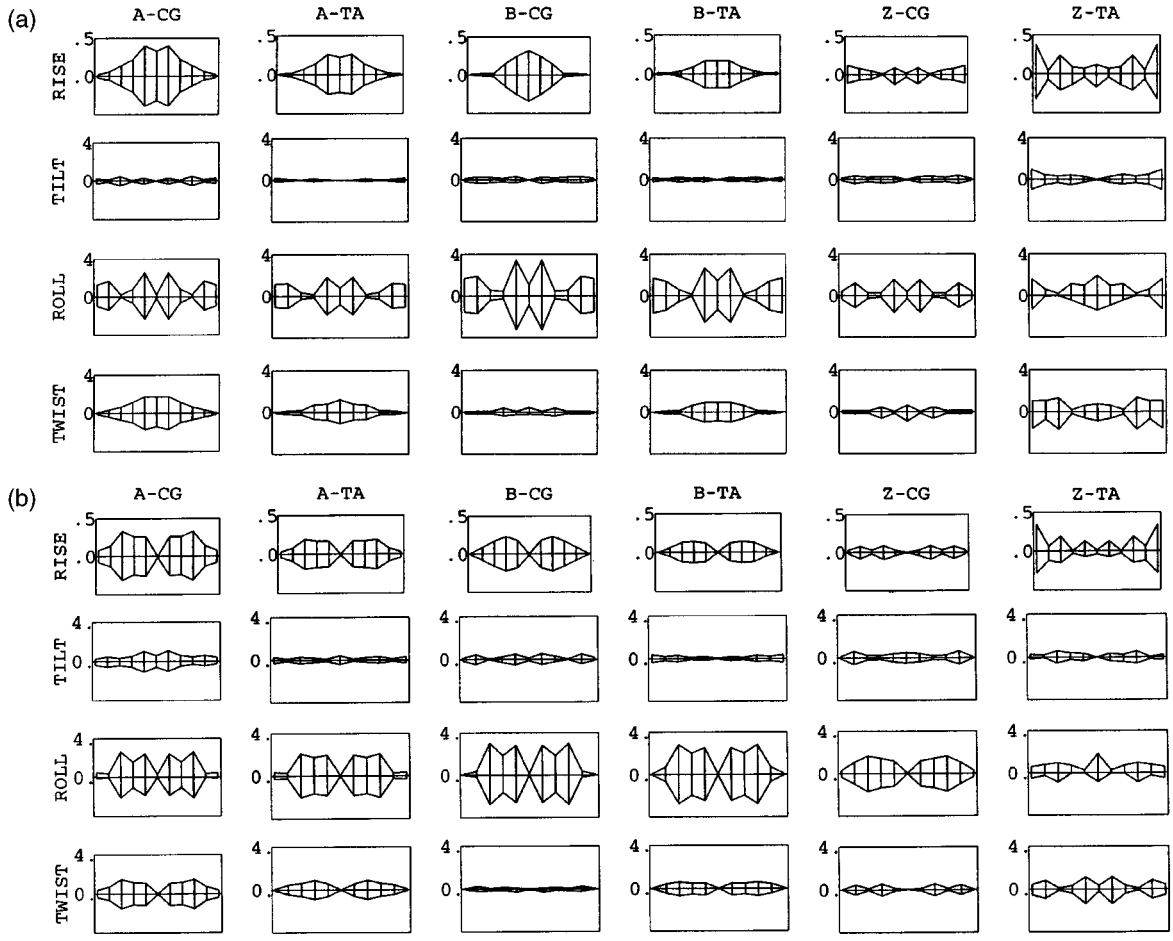


**FIGURE 6.** Amplitudes of rise (Å), tilt (°), roll (°), and twist (°) variations along the dodecamers studied for the two lowest frequency bending modes. (a) First bending mode. (b) Second bending mode.

twist, and roll for the first bending mode of B-CG in Figure 7. Movements of twist and roll are clearly anticorrelated. Similar behavior has recently been observed in other studies, including an analysis of crystallized DNA–protein complexes[29] and modeling of DNA supercoiling.[30] These static studies showed that high positive roll is coupled to low values of twist and it is interesting to see the same relationship also appearing in dynamic behavior.

## THERMODYNAMIC PROPERTIES

In Table IV we give the values of vibrational free energy and entropy per base pair calculated for the three conformers of the two sequences. The magnitude of the vibration free energy follows the order B > Z > A. The change in internal stability is therefore less for the passage B → Z than for B → A in agreement with earlier calculations for the CG alternating sequence.[31] Although environmental effects involving solvent and counterion rearrangements are clearly important in such transitions, it is interesting to note that the calculated free energy change of 0.4 to 0.5 kcal/mol per base pair for B–Z transition is rather close to the experimental data.[32]

## INFLUENCE OF FRAGMENT LENGTH

To study the effect of oligomer length, we performed NM calculations on five B-DNA oligomers with the alternating CG sequence with lengths ranging from 6 to 18 base pairs. Figure 8 shows the variation of the range and frequency of the first two bending modes and the global twisting mode as a function of length. It can be seen that the frequencies move to lower values with an increasing number of base pairs in agreement with classi-

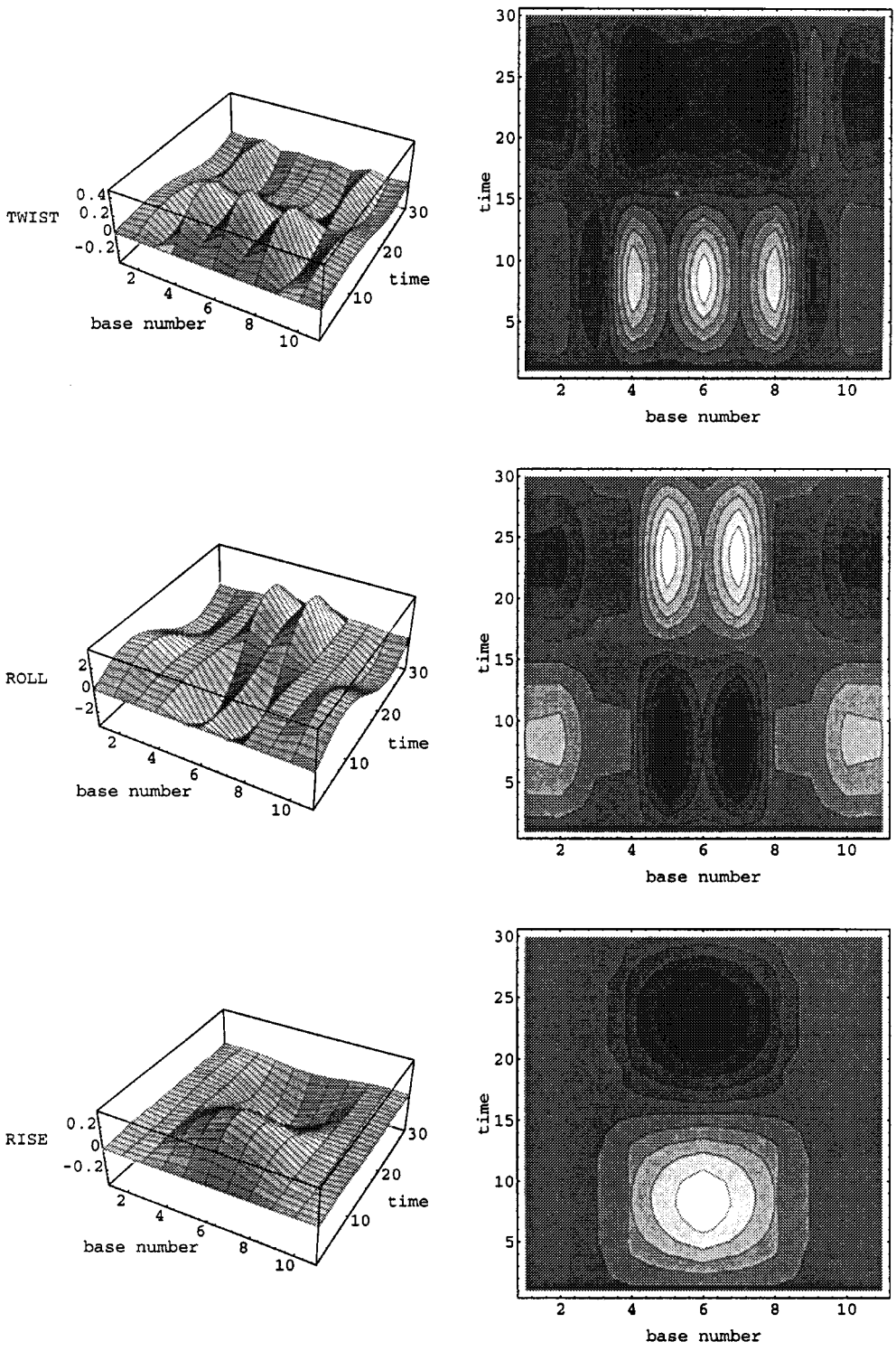


**FIGURE 7.** Time evolution of twist (°), roll (°), and rise (Å) for the first bending mode of B-CG.

**TABLE IV.**
**Contribution of Low Frequency Modes to Vibrational Free Energy and Entropy ($F_{vib}$ in Kilocalories per Mole per Base Pair and $S_{vib}$ in Calories per Mole per Degrees Kelvin per Base Pair).**

| | A-DNA | | B-DNA | | Z-DNA | | B-A | | B-Z | |
|---|---|---|---|---|---|---|---|---|---|---|
| | CG | TA | CG | TA | CG | TA | CG | TA | CG | TA |
| $F_{vib}$ | −3.53 | −4.31 | −4.49 | −5.48 | −4.05 | −4.96 | −0.96 | −1.17 | −0.44 | −0.52 |
| $S_{vib}$ | 64.8 | 69.3 | 68.0 | 71.9 | 65.6 | 70.4 | 3.2 | 2.6 | 2.4 | 1.5 |

cal mechanical principles (similar results were observed experimental by Peticolas[33] in the case of polypeptides in an α-helical conformation). The amplitude of the movements also increase with length, in the same way as for a solid rod. However, there is no simple relationship between these two parameters.

## COMPARISON WITH EXPERIMENTAL DATA

Experimental data concerning low frequency vibrations are scarce and difficult to interpret because of lack of conformational data. As an illustration of this point, Weidlich et al.[34] found at least five Raman active modes under 200 $cm^{-1}$ in the spectra of both A- and B-DNA crystals. Four of these modes are believed to be intrahelical modes, whereas the mode at 25 $cm^{-1}$ is thought to involve interactions between neighboring oligomers in the crystal. It is difficult to compare such data with a normal mode analysis which yields several hundred vibrations of frequency under 200 $cm^{-1}$. In contrast, recent NMR and fluorescence polarization anisotropy (FPA) studies on nucleic acids fragments succeeded in quantifying some dynamic features, such as the amplitude of local angular motions.[35,36] These studies indicate vibrational movements with amplitudes of about 10° in bases and 20° in sugar rings. In the phosphodiester backbone, the amplitudes are found to be 2° to 6° larger

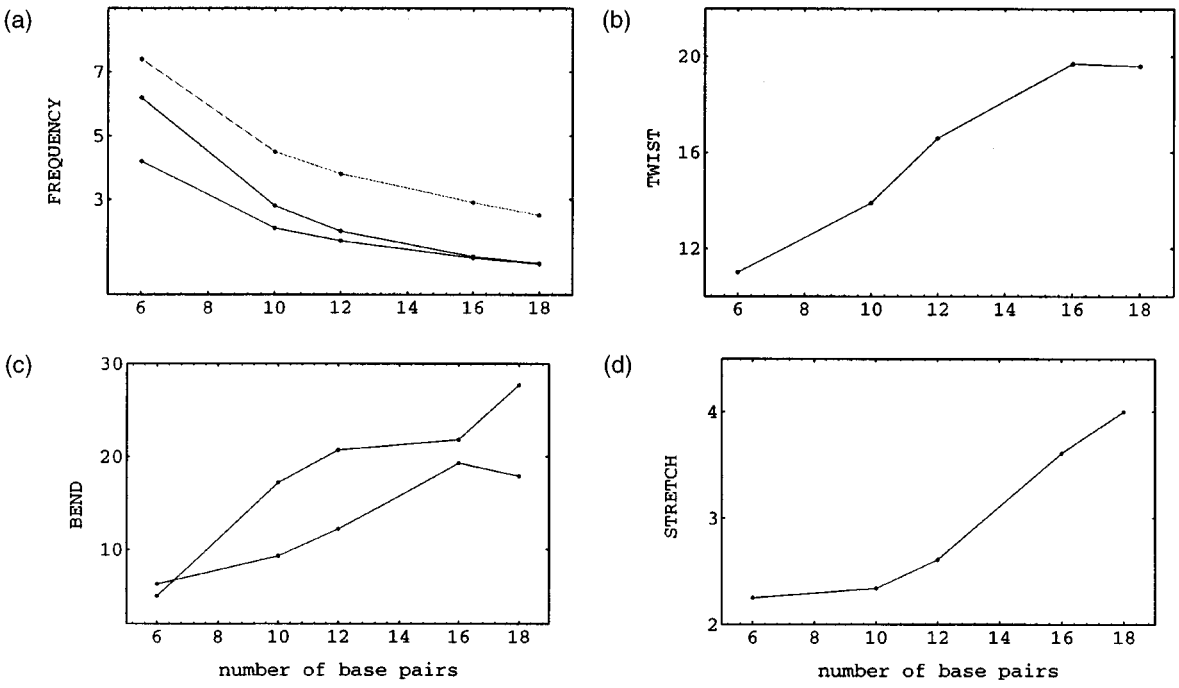


**FIGURE 8.** Dependence of the lowest frequency modes on the oligomer length. (a) Frequency ($cm^{-1}$) of the two lowest global bending modes (black line) and of the global twisting mode (gray line). (b) Amplitude (°) of global twisting mode. (c) Amplitude (°) of the two lowest global bending modes. (d) Amplitude (Å) of the global stretching mode.

than amplitude in the sugars.[36] The results of our analysis are in reasonable accord with these data because we also observe mobilities in the order: bases < sugars < backbone. Calculated backbone dihedral variations indeed reach roughly 20° with characteristic times up to roughly 1 ps.

Concerning the global deformations of DNA, an encouraging agreement is found with FPA results on short restriction fragments.[37] By extrapolating these results, Kintanar et al.[38] have calculated that an oligonucleotide of 12 base pairs should have a range of twisting movement around 18°, which is very close to our calculated global twisting variation ($\sim 17°$). Another extrapolation of FPA results[39] yields an order of magnitude of the characteristic times of global bending motions (assuming, as have the authors of this study, that the relaxation time for the longest bending mode will vary as the fourth power of length). This value of 9 to 10 ps, corresponding to a frequency of roughly 3 to 4 $cm^{-1}$ is once again close to the value we have obtained.

## COMPARISON WITH EARLIER NM STUDIES AND TESTS OF ANHARMONICITY

To compare the normal modes calculated in the DAS (and with our FLEX force field) with earlier calculations we have calculated the normal modes of a short B-DNA oligomer $d(CG)_3 \cdot d(CG)_3$ used by other investigators for Cartesian coordinate space calculations with the AMBER and CHARMM force fields.[40-42] In Table V, a comparison is given of the 10 lowest frequencies obtained by these different methods. Despite the methodological differences these 10 lowest frequencies are roughly in the same range. Overall, our values fall between the low frequencies of the AMBER force field (slightly increased by the united atom model) and the high frequencies obtained with CHARMM.

To characterize the anharmonicity of the calculated normal modes, we used the procedure proposed by Noguti and Go[43] and Garcia.[31] This involves calculating the exact potential energy for the conformational changes in the direction of each mode and fitting these curves with a fourth order polynomial ($Bx^2 + Cx^3 + Dx^4$). As expected, the lowest frequencies show an anharmonic character. This can be quantified by the ratios $C/B$ and $D/B$ which characterize the third and fourth order contributions. According to the classification proposed by Noguti and Go[43] almost one third of the lowest frequency modes should be classified as anharmonic. This implies that the lowest frequency normal modes underestimate the associated deformation energies and that their calculated vibrational amplitudes should be taken as an upper limit and that the associated movements may in fact be more anisotropic. This anharmonic character could also be increased by taking in account solvent damping.

Despite these restrictions it appears that the dihedral angle space normal modes we have calculated for DNA can be a useful guide to the mechanisms of dynamic behavior and show reasonable correlation with available experimental data.

**TABLE V.**
**The 10 Lowest Frequencies of Vibration (in Centimeter $^{-1}$) Obtained by Four Different Normal Mode Calculations on d(CGCGCG).**

| Our study[a] | Garcia et al.[b] | Kottalam et al.[c] | Tidor et al.[d] |
|---|---|---|---|
| 4.2 | 2.5 | 4.1 | 9.7 |
| 6.2 | 3.5 | 4.7 | 10.5 |
| 7.4 | 3.9 | 4.8 | 12.5 |
| 9.7 | 5.3 | 7.4 | 15.8 |
| 11.0 | 5.6 | 7.8 | 16.1 |
| 13.2 | 6.3 | 8.2 | 19.2 |
| 14.1 | 7.4 | 9.2 | 21.0 |
| 14.9 | 8.2 | 9.3 | 22.1 |
| 15.2 | 9.6 | 10.2 | 24.1 |
| 16.1 | 9.9 | 10.6 | 25.8 |

[a] FLEX with all-atom model.
[b] AMBER with all-atom model.[28]
[c] AMBER with united atom model.[29]
[d] CHARMM with united atom model.[27]

## Conclusion

This study describes the implementation of normal mode calculations in dihedral angle space for nucleic acids taking into account sugar ring flexibility. It also presents means of analyzing the resulting movements in terms of helicoidal and backbone parameters. This combined approach allows the dynamic behavior of nucleic acids fragments to be calculated rapidly and understood in an objective and automated way. It opens the route to studying sequence and conformational effects on nucleic acid dynamics for both isolated molecules and protein–nucleic acid complexes.

A first application of this technique has shown that the allomorphic form of DNA has the largest influence on flexibility of the sugar–phosphate backbone. Vibrational amplitudes are generally largest in B-DNA and smallest in the A-DNA, with Z-DNA representing the intermediate case. In contrast, the amplitudes of helicoidal parameters are much more dependent on the base sequence.

An analysis of the lowest frequency modes has shown that they correspond to collective movements involving global deformation of the double helix and, in many aspects, resemble standing waves. They occur with characteristic times in the range of 1 to 10 ps and can be of mixed character, the major bending mode being at the same time the major stretching mode. Dynamic correlations between helicoidal parameters are found that correlate with similar coupling in static, deformed DNA structures (notably, roll, twist, and rise during DNA bending).

## Acknowledgments

The authors thank the Association for International Cancer Research (St. Andrews, U.K.) for their generous funding of this research.

## Appendix A: Calculation of Potential Energy Matrix F

Before writing the mathematical expressions of the second derivatives of the potential energy, we define the notations we employ:

- The three atoms involved in hydrogen bonds will be denoted $M—H \cdots B$, where $B$ is the hydrogen bond acceptor, $H$ the proton donor, and $M$ the atom linked to $H$. The vector $\mathbf{p}$ is a unit vector in the direction of the M—H bond. $\theta$ is the angle between $\mathbf{p}$ and the hydrogen bond $H \cdots B$.
- We distinguish two kinds of derivatives: $d(\ )/dq_i$ will represent a derivative with respect to any variable of our model (dependent or independent) without taking into account the eventual variations in the other variables, whereas $\partial(\ )/\partial q_i$ represents a derivative with respect to the independent variables:

$$\frac{\partial(\ )}{\partial q_{\text{ind}}} = \frac{d(\ )}{dq_{\text{ind}}} + \sum_{q_{\text{dep}}} \frac{d(\ )}{dq_{\text{dep}}} \cdot \frac{dq_{\text{dep}}}{dq_{\text{ind}}}$$

From the equation of the potential energy, the first derivatives with respect to independent variables can be written in the form:

$$\frac{\partial Ep}{\partial q_i} = T_i + \sum_{a \in \Omega_i} \Gamma_{ai}(\mathbf{F}_a \times \mathbf{r}_{ai}) \cdot \mathbf{e}_i$$

where $\Omega_i$ is the subset of atoms which move when the variable $q_i$ change. $\Gamma_{ai}$ is normally equal to 1, but becomes 0.5 for the substituents of an endocyclic valence angle.[15] $\mathbf{e}_i$ is a unit vector parallel to the axis of rotation. $\mathbf{r}_{ai}$ is a vector joining the atom $a$ to a point on the rotation axis. $T_i$ is the magnitude of the torque produced by the derivative of the energetic term for the valence angle or the dihedral angle which moves:

$$T_i = \mp \frac{V_{i(\text{tor})}}{2} n_i \sin(n_i \tau_i) \text{ or}$$

$$T_i = 2V_{i(\text{val})}(\sigma_i - \sigma_i^0)$$

The atomic forces are given by the formula:

$$\mathbf{F}_a = \sum_b \left[ \frac{q_a q_b}{r_{ab}^2} \left( \frac{1}{\varepsilon r_{ab}} + \frac{\varepsilon'}{\varepsilon^2} \right) - \left( 6\frac{A_{ab}}{r_{ab}^8} - 12\frac{B_{ab}}{r_{ab}^{14}} \right) \right] \mathbf{r}_{ab} + \mathbf{F}_{a(\text{pen})} + \mathbf{F}_{a(\text{cyc})} + \mathbf{F}_{a(\text{hyd})}$$

$\mathbf{F}_{a(\text{pen})}$ only exists if atom $a$ is one of the two atoms between which we apply the ring closure constraint:

$$\mathbf{F}_{a(\text{pen})} = \pm 2K(r_{12} - r_{12}^0)\frac{\mathbf{r}_{12}}{r_{12}}$$

The presence of $\mathbf{F}_{a(\text{cyc})}$ in the equation of atomic force is due to the derivatives of $Ep$ with respect to dependant angles. $\mathbf{F}_{a(\text{cyc})}$ is non-null only in the case where atom $a$ enters into the definition of a dependent endocyclic angle:

$$\mathbf{F}_{a(\text{cyc})} = -\sum_{q_{\text{dep}} \supset a} \frac{dEp}{dq_{\text{dep}}}\mathbf{grad}_a(q_{\text{dep}})$$

The summation is over all dependent angles involving atom $a$. Finally, the last term, $\mathbf{F}_{a(\text{hyd})}$, concerns the three atoms involved in hydrogen bonding.

- If $a$ is the acceptor ($B$) then:

$$\mathbf{F}_{a(\text{hyd})} = \sum_{b=H}\left[\begin{array}{l}\left(\dfrac{A_{ab}^{HB} - A_{ab}}{r_{ab}^7} - \dfrac{B_{ab}^{HB} - B_{ab}}{r_{ab}^{13}}\right)\mathbf{p} \\ -\cos\theta\left(7\dfrac{A_{ab}^{HB} - A_{ab}}{r_{ab}^8} - 13\dfrac{B_{ab}^{HB} - B_{ab}}{r_{ab}^{14}}\right)\mathbf{r}_{ab}\end{array}\right]$$

- If $a$ is the proton then:

$$\mathbf{F}_{a(\text{hyd})} = \sum_{b=B}\left[\begin{array}{l}-\left(\dfrac{A_{ab}^{HB} - A_{ab}}{r_{ab}^7} - \dfrac{B_{ab}^{HB} - B_{ab}}{r_{ab}^{13}}\right)\mathbf{p} \\ -\cos\theta\left(7\dfrac{A_{ab}^{HB} - A_{ab}}{r_{ab}^8} - 13\dfrac{B_{ab}^{HB} - B_{ab}}{r_{ab}^{14}}\right)\mathbf{r}_{ab} \\ -\left(\dfrac{A_{ab}^{HB} - A_{ab}}{r_{ab}^7} - \dfrac{B_{ab}^{HB} - B_{ab}}{r_{ab}^{13}}\right)\dfrac{\mathbf{r}_{ab}}{r_{MH}}\end{array}\right]$$

- If $a$ is the atom linked to the proton donor ($M$) then:

$$\mathbf{F}_{a(\text{hyd})} = \sum_{b=B}\sum_{c=H} -\left(\frac{A_{bc}^{HB} - A_{bc}}{\mathbf{r}_{bc}^7} - \frac{B_{bc}^{HB} - B_{bc}}{r_{bc}^{13}}\right)\frac{\mathbf{r}_{bc}}{r_{MH}}$$

From all these equations, we can calculate the second derivatives of the potential energy with respect to two independent variables as follows:

$$\frac{\partial^2 Ep}{\partial q_j\,\partial q_i} = \frac{\partial T_i}{\partial q_j} + \sum_{a\in\Omega_i}\Gamma_{ai}\left[\left(\frac{\partial \mathbf{F}_a}{\partial q_j}\times\mathbf{r}_{ai} + \mathbf{F}_a\times\frac{\partial \mathbf{r}_{ai}}{\partial q_j}\right)\cdot\mathbf{e}_i + (\mathbf{F}_a\times\mathbf{r}_{ai})\cdot\frac{\partial \mathbf{e}_i}{\partial q_j}\right]$$

where:

$$\frac{\partial T_i}{\partial q_j} = \mp\delta_{ij}\frac{V_{i(\text{tor})}}{2}n_i^2\cos(n_i\tau_i) \text{ or } \frac{\partial T_i}{\partial q_j} = \delta_{ij}2V_{i(\text{val})}$$

We can recall here that the derivative of a vector $\mathbf{r}_a$ with respect to any angle $q_i$ is given by:

$$\frac{d\mathbf{r}_a}{dq_i} = \Gamma_{ai}(\mathbf{e}_i\times\mathbf{r}_{ai})$$

To simplify the expressions for the derivatives of the atomic forces, we introduce the following notations:

$$C_{ab} = \frac{q_a q_b}{r_{ab}^3}\left(\frac{\varepsilon''}{\varepsilon^2} - \frac{2\varepsilon'^2}{\varepsilon^3} - \frac{3\varepsilon'}{\varepsilon^2 r_{ab}} - \frac{3}{\varepsilon r_{ab}^2}\right) + \left(48\frac{A_{ab}}{r_{ab}^{10}} - 168\frac{B_{ab}}{r_{ab}^{16}}\right)$$

$$D_{ab} = \left(63\frac{A_{ab}^{HB} - A_{ab}}{r_{ab}^{10}} - 195\frac{B_{ab}^{HB} - B_{ab}}{r_{ab}^{16}}\right)$$

$$E_{ab} = \frac{q_a q_b}{r_{ab}^2}\left(\frac{1}{\varepsilon r_{ab}} + \frac{\varepsilon''}{\varepsilon^2}\right) - \left(6\frac{A_{ab}}{r_{ab}^8} - 12\frac{B_{ab}}{r_{ab}^{14}}\right)$$

$$F_{ab} = \left(7\frac{A_{ab}^{HB} - A_{ab}}{r_{ab}^9} - 13\frac{B_{ab}^{HB} - B_{ab}}{r_{ab}^{13}}\right)$$

$$G_{ab} = \left(\frac{A_{ab}^{HB} - A_{ab}}{r_{ab}^7} - \frac{B_{ab}^{HB} - B_{ab}}{r_{ab}^{13}}\right)$$

With these notations, we can write that:

$$\frac{d\mathbf{F}_a}{dq_j} = \sum_b\left[C_{ab}\left(\mathbf{r}_{ab}\cdot\frac{d\mathbf{r}_{ab}}{dq_j}\right)\mathbf{r}_{ab} + E_{ab}\frac{d\mathbf{r}_{ab}}{dq_j}\right] + \frac{d\mathbf{F}_{a(\text{pen})}}{dq_j} + \frac{d\mathbf{F}_{a(\text{eyc})}}{dq_j} + \frac{d\mathbf{F}_{a(\text{hyd})}}{dq_j}$$

where:

$$\frac{d\mathbf{F}_{a(\text{pen})}}{dq_j} = \pm 2\frac{K}{r_{12}}\left[\frac{r_{12}^0}{r_{12}^2}\left(\mathbf{r}_{12}\cdot\frac{d\mathbf{r}_{12}}{dq_j}\right)\mathbf{r}_{12} + (r_{12}-r_{12}^0)\frac{d\mathbf{r}_{12}}{dq_j}\right]$$

$$\frac{d\mathbf{F}_{a(\text{cyc})}}{dq_j} = \sum_{q_{\text{dep}}\supset a}\left\{\frac{d^2Ep}{dq_j\,dq_{\text{dep}}}\mathbf{grad}_{\text{a}}(\text{q}_{dep}) + \frac{\text{dEp}}{\text{dq}_{dep}}\frac{\text{d}\left[\mathbf{grad}_{\text{a}}(\text{q}_{dep})\right]}{\text{dq}_{\text{j}}}\right\}$$

- If $a = B$:

$$\frac{d\mathbf{F}_{a(\text{hyd})}}{dq_j} = \sum_{b=H}\left[\begin{array}{l} -F_{ab}\left(\mathbf{r}_{ab}\cdot\dfrac{d\mathbf{r}_{ab}}{dq_j}\right)\mathbf{p} + G_{ab}\dfrac{d\mathbf{p}}{dq_j} \\ +\cos\theta D_{ab}\left(\mathbf{r}_{ab}\cdot\dfrac{d\mathbf{r}_{ab}}{dq_j}\right)\mathbf{r}_{ab} - \cos\theta F_{ab}r_{ab}\dfrac{d\mathbf{r}_{ab}}{dq_j} \\ -F_{ab}\left(\dfrac{d\mathbf{p}}{dq_j}\cdot\mathbf{r}_{ab} + \mathbf{p}\cdot\dfrac{d\mathbf{r}_{ab}}{dq_j}\right)\mathbf{r}_{ab} \end{array}\right]$$

- If $a = H$:

$$\frac{d\mathbf{F}_{a(\text{hyd})}}{dq_j} = \sum_{b=N}\left[\begin{array}{l} F_{ab}\left(\mathbf{r}_{ab}\cdot\dfrac{d\mathbf{r}_{ab}}{dq_j}\right)\mathbf{p} - G_{ab}\dfrac{d\mathbf{p}}{dq_j} + \cos\theta D_{ab}\left(\mathbf{r}_{ab}\cdot\dfrac{d\mathbf{r}_{ab}}{dq_j}\right)\mathbf{r}_{ab} \\ -\cos\theta F_{ab}r_{ab}\dfrac{d\mathbf{r}_{ab}}{dq_j} - F_{ab}\left(\dfrac{d\mathbf{p}}{dq_j}\cdot\mathbf{r}_{ab} + \mathbf{p}\cdot\dfrac{d\mathbf{r}_{ab}}{dq_j}\right)\mathbf{r}_{ab} \\ +F_{ab}\left(\mathbf{r}_{ab}\cdot\dfrac{d\mathbf{r}_{ab}}{dq_j}\right)\dfrac{\mathbf{r}_{ab}}{r_{MH}} - G_{ab}\dfrac{1}{r_{MH}}\dfrac{d\mathbf{r}_{ab}}{dq_j} \end{array}\right]$$

- If $a = M$:

$$\frac{d\mathbf{F}_{a(\text{hyd})}}{dq_j} = \sum_{b=B}\sum_{c=H}\left[F_{bc}\left(\mathbf{r}_{bc}\cdot\frac{d\mathbf{r}_{bc}}{dq_j}\right)\frac{\mathbf{r}_{bc}}{r_{MH}} - G_{bc}\frac{1}{r_{MH}}\frac{d\mathbf{r}_{ab}}{dq_j}\right]$$

## Appendix B: Calculation of Kinetic Matrix H

For all angles that are not endocyclic valence angles, we use the Noguti and Go[19] approach to calculate the derivative of atomic positions with elimination of overall translation and rotation. Thus, for any atom $b$ belonging to the first body (see Fig. 3) characterized by a mass, $M_1$, and a center of mass, $G_1$, and for any atom, $c$, of the second body characterized by $M_2$ and $G_2$, the calculations yield:

$$\frac{d\mathbf{r}_b}{dq_i} = \mathbf{e}_i\times\left[\left(1-\frac{M_1}{M_1+M_2}\right)\mathbf{r}_i^0 + \frac{M_1}{M_1+M_2}\mathbf{G}_1^0\right] - \left\{\mathbf{I}^{-1}\left[M_1\mathbf{G}_1^0\times(\mathbf{e}_i\times\mathbf{r}_i^0) + (\mathbf{I}-\mathbf{I}_1)\mathbf{e}_i\right]\right\}\times\mathbf{r}_b^0$$

$$\frac{d\mathbf{r}_c}{dq_i} = -\mathbf{e}_i\times\left[\left(1-\frac{M_2}{M_1+M_2}\right)\mathbf{r}_i^0 + \frac{M_2}{M_1+M2}\mathbf{G}_2^0\right] + \left\{\mathbf{I}^{-1}\left[M_2\mathbf{G}_2^0\times(\mathbf{e}_i\times\mathbf{r}_i^0) + (\mathbf{I}-\mathbf{I}_2)\mathbf{e}_i\right]\right\}\times\mathbf{r}_c^0$$

The two matrices, $\mathbf{I}_1$ and $\mathbf{I}_2$, are $3 \times 3$ matrices whose elements are:

$$I_{ij} = \sum_a m_a \left[ \left( \delta_{ij} \sum_k r_{ak} \right) - r_{ai} r_{aj} \right]$$

For $\mathbf{I}_1$ the summation is over all atoms of the first body, and for $\mathbf{I}_2$ the summation is over all atoms of the second body. $\mathbf{I} = \mathbf{I}_1 + \mathbf{I}_2$. As mentioned in the Methodology section, the derivatives of atomic positions with respect to the endocyclic valence angles are calculated by regarding the molecule as three rigid bodies (see Fig. 4). In this case, with the same notations as those just expressed, we define two vectors, **W** and **G**, as:

$$\mathbf{W} = \mathbf{I}^{-1} \left[ \left( M_2 \mathbf{G}_2^0 + \frac{1}{2} M_3 \mathbf{G}_3^0 \right) \times (\mathbf{e}_i \times \mathbf{r}_i^0) - \left( \mathbf{I}_2 + \frac{1}{2} \mathbf{I}_3 \right) \mathbf{e}_i \right]$$

$$\mathbf{G} = \frac{M_2}{M} \mathbf{G}_2^0 + \frac{1}{2} \frac{M_3}{M} \mathbf{G}_3^0 - \frac{1}{2} \left( \frac{M_2}{M} + \frac{1}{2} \frac{M_3}{M} \right) \mathbf{r}_i^0$$

and, by using these two vectors, we can formulate the final results of the calculations:

$$\frac{d\mathbf{r}_b}{dq_i} = \mathbf{W} \times \mathbf{r}_b^0 - \mathbf{e}_i \times \mathbf{G}$$

$$\frac{d\mathbf{r}_c}{dq_i} = \mathbf{W} \times \mathbf{r}_c^0 - \mathbf{e}_i \times (\mathbf{G} + \mathbf{r}_i^0 - \mathbf{r}_c^0)$$

$$\frac{d\mathbf{r}_d}{dq_i} = \mathbf{W} \times \mathbf{r}_d^0 - \mathbf{e}_i \times \left( \mathbf{G} + \mathbf{r}_i^0 - \frac{1}{2} \mathbf{r}_d^0 \right)$$

## References


1. G. Weber, Adv. Prot. Chem., **29,** 1 (1975).
2. P. G. Debrunner and H. Frauenfelder, Annu. Rev. Phys. Chem., **33,** 283 (1982).
3. J. A. McCammon, Rep. Prog. Phys., **47,** 1 (1984).
4. R. G. Brown, S. C. Erfurth, E. W. Small, and W. L. Peticolas, Proc. Natl. Acad. Sci. USA, **69,** 1467 (1972).
5. W. L. Peticolas, Meth. Enzymol., **61,** 425 (1979).
6. H. D. Bartunik, P. Jolles, J. Berthou, and A. J. Dianoux, Biopolymers, **21,** 43 (1982).
7. J. C. Smith, Quart. Rev. Biophys., **24,** 227 (1991).
8. D. Perahia and L. Mouawad, Comput. Chem., **19,** 241 (1995).
9. P. Durand, G. Trinquier, and Y. H. Sanejouand, Biopolymers, **34,** 759 (1994).
10. S. Sunada and N. Go, J. Comput. Chem., **11,** 328 (1995).
11. A. Kitao and N. Go, J. Comput. Chem., **12,** 359 (1991).
12. A. Kitao, S. Hayward, and N. Go, Biophys. Chem., **52,** 107 (1994).
13. R. Lavery and H. Sklenar, J. Biomol. Struct. Dynam., **6,** 655 (1989).
14. H. Goldstein, Classical Mechanics, Addison-Wesley, Reading, MA, 1950.
15. R. Lavery, I. Parker, and J. Kendrick, J. Biomol. Struct. Dynam., **4,** 443 (1986).
16. R. Lavery, H. Sklenar, K. Zakrzewska, and B. Pullman, J. Biomol. Struct. Dynam., **3,** 989 (1986).
17. R. Lavery, Structure and expression, Vol. 3: DNA Bending and Curvature, W. K. Olson, R. H. Sarma, M. H. Sarma, and M. Sundaralingam, Eds., Adenine Press, New York, 1988, p. 191.
18. R. Lavery, K. Zakrzewska, and H. Sklenar, Comput. Phys. Commun., **91,** 135 (1995).
19. T. Noguti and N. Go, J. Phys. Soc. Jpn., **52,** 3283 (1983).
20. C. Eckart, Phys. Rev., **47,** 552 (1935).
21. M. Levitt, C. Sander, and P. S. Stern, J. Mol. Biol., **181,** 423 (1985).
22. R. Lavery and B. Hartman, Biophys. Chem., **60,** 33 (1994).
23. H. Gabb, R. Lavery, and C. Prevost, J. Comput. Chem., **16,** 667 (1995).
24. S. T. Rao, E. Westhof, and M. Sundaralingam, Acta Cryst., **37,** 421 (1981).
25. G. G. Privé, U. Heinemann, S. Chadrasekharan, L. S. Kan, M. L. Kopka, and R. E. Dickerson, Science, **238,** 498 (1987).
26. A. D. DiGabriele, M. R. Sanderson, and T. A. Steitz, Proc. Natl. Acad. Sci. USA, **16,** 1816 (1989).
27. V. B. Zhurkin, Y. P. Lysov, and V. I. Ivanov, Nucl. Acids Res., **6,** 1081 (1979).
28. W. K. Olson, A. R. Srinivasan, R. C. Maroun, R. Torres, and W. Clark, Unusual DNA Structures, R. D. Wells, and S. C. Harvey, Eds., Springer, New York, 1989, p. 207.
29. M. Suzuki and N. Yagi, Nucl. Acids Res., **23,** 2083 (1995).
30. S. R. Sanghani, K. Zakrzewska, S. C. Harvey, and R. Lavery, Nucl. Acids Res. (in press).
31. A. E. Garcia, Computation of Biomolecular Structures, D. M. Soumpasis and T. M. Jovin, Eds., Springer, New York, 1992, p. 165.
32. T. M. Jovin, D. M. Soumpasis, and L. P. McIntosh, Annu. Rev. Phys. Chem., **38,** 521 (1987).
33. W. L. Peticolas, Biopolymers, **18,** 747 (1979).
34. T. Weidlich, S. M. Linsay, Q. Rui, A. Rupprecht, W. L. Peticolas, and G. A. Thomas, J. Biomol. Struct. Dynam., **8,** 139 (1990).
35. S. Nuutero, B. S. Fujimoto, P. F. Flynn, B. R. Reid, and J. M. Schurr, Biopolymers, **34,** 463 (1994).
36. T. M. Alam, J. Orban, and G. P. Drobny, Biochemistry, **30,** 9229 (1991).
37. J. H. Shibata, B. S. Fujimoto, and J. M. Schurr, Biopolymers, **24,** 1909 (1985).
38. A. Kintanar, W. C. Huang, D. C. Schindele, D. E. Wemmer, and G. Drobny, Biochemistry, **28,** 282 (1989).
39. P. Wu, B. S. Fujimoto, and J. M. Schurr, Biopolymers, **26,** 1463 (1987).
40. B. Tidor, K. K. Irikura, B. P. Brooks, and M. Karplus, J. Biomol. Struct. Dynam., **1,** 231 (1983).
41. A. E. Garcia and D. M. Soumpasis, Proc. Natl. Acad. Sci. USA, **86,** 3160 (1972).
42. J. Kottalam and D. A. Case, Biopolymers, **29,** 1409 (1990).
43. T. Noguti and N. Go, Nature, **296,** 776 (1982).